NORTHCENTRAL UNIVERSITY
ASSIGNMENT COVER SHEET

Student: **Michael Higley-Vance**

THIS FORM MUST BE COMPLETELY FILLED IN

Follow these procedures: If requested by your instructor, please include an assignment cover sheet. This will become the first page of your assignment. In addition, your assignment header should include your last name, first initial, course code, dash, and assignment number. This should be left justified, with the page number right justified. For example:

DoeJXXX0000-1	1

Save a copy of your assignments: You may need to re-submit an assignment at your instructor's request. Make sure you save your files in accessible location.

Academic integrity: All work submitted in each course must be your own original work. This includes all assignments, exams, term papers, and other projects required by your instructor. Knowingly submitting another person's work as your own, without properly citing the source of the work, is considered plagiarism. This will result in an unsatisfactory grade for the work submitted or for the entire course. It may also result in academic dismissal from the University.

EDU7006-8	**Dr. Rebecca Watts**
Quantitative Research Design	**Activity #2a: Exploring Inferential Statistics**

Comments: I'm just now finishing with the first assignment in this two-part assignment. I can't believe this one was only worth 5pts. I'm hoping to finish the second part by Sunday afternoon.

Faculty Use Only

~~<Faculty comments here>~~Michael, I do realize the difficulty of this assignment. It is probably the most difficult of all assignments. I appreciate your hard work. As you have probably noticed, the key to calculating the test statistics is to locate the correct formula and insert the correct values into that formula. Sometimes, you must calculate one value from another value. Thus, if you miscalculate one value, it can affect your calculation of other values. In your paper, you use the correct formulas and you approaching the assignments correctly. Your are following the process, but you are not calculating the values correctly.

I might suggest that you try to do your calculations in Excel. Excel will calculate the statistics for you. As I said, you are reasoning correctly in solving the problems, but your calculations are not correct. You are stating the hypotheses correctly and you are using the critical tables to determine the critical values. You are making calculations mistakes and this affects all of your calculations.

I want you to write the following on a note care:

If the absolute value of the calculated statistic is greater than the critical table value of the statistic, then you REJECT the null hypothesis.

If the absolute value of the calculated statistic is less than the critical table value of the statistic (at the respective alpha level), then you DO NOT REJECT the null hypothesis.

It is most important that you apply the rule correctly. It is important that you calculate the test statistic correctly.

I am attaching a file here that shows the correct calculations for the statistics in assignment 2. I ask that you not share this with other students. Review the calculations and see if you can determine where you made the calculation errors. I also encourage you to call me if you have questions or feel overwhelmed.

Answers to Activity 2.docx

Score = 82

<Faculty Name> <Grade Earned> <Writing Score> <Date Graded>

Numerical Points	Letter Grade	Descriptor	Explanation

100 - 94	**A**	**Excellent**	Completes **all** required parts of the assignment, demonstrates **deep** understanding of materials, uses **very** clear and effective expression appropriate to scholarly writing, and has **very few or no** errors in grammar, mechanics, and APA formatting.
93-90	**A-**		
89-87	**B+**	**Good**	Completes **all or most** required parts of the assignment, demonstrates **good** understanding of readings, uses **mostly** clear and effective expression appropriate to scholarly writing, and has **few** errors in grammar, mechanics, and APA formatting.
86-83	**B**		
82-80	**B-**	**Fair**	Completes **most** required parts of the assignment, demonstrates **some** understanding of readings, and writing is **somewhat** clear, effective , and scholarly, and has **some** errors in grammar, mechanics, and APA formatting.
79-77	**C+**		
76-73	**C**	**Poor**	Completes **some** required parts of the assignment, demonstrates **some** understanding of readings, and writing is **difficult to understand** and **unscholarly** and has **several** errors in grammar, mechanics, and APA formatting.
72-0	**F**	**Unacceptable**	Completes **few** required parts of the assignment, demonstrates **little** understanding of readings, and writing is **difficult to understand** and **unscholarly** and has **many** errors in grammar, mechanics, and APA formatting.

The Logic of Experimental Design

1. Jackson even-numbered Chapter Exercises (pp. 220-221).
   - Question 2. The producers of a new toothpaste claim that it prevents more cavities than other brands of toothpaste. A random sample of 60 people used the new toothpaste for 6 months. The mean number of cavities at their next checkup is 1.5. In the general population, the mean number of cavities at a 6-month checkup is 1.73 (σ =1.12).

     a.) Is this a one- or two-tailed test? one-tailed test
     b.) What are H_0 and H_a for this study? H_0: $\mu \geq 1.73$ and H_a: $\mu < 1.73$
     c.) Compute z_{obt}. Z obtained is z_{obt} equals the average of the sample (Xbar)minus the average of the population (μ), divided by the standard deviation of the sample (σ), divided be the square root of N ($\sqrt{n}$).

     Example: $z_{obt} = (\bar{x} - \mu) \div (\sigma \div (\sqrt{n}))$

     $n = 60$, $\bar{x} = 1.5$, $\mu = 1.73$, $\sigma = 1.12$
     $(1.5 - 1.73) \div (1.12 \div (\sqrt{60}))$
     $-.23 \div (1.12 \div 7.75)$
     $-.23 \div 0.1445 = -1.59$ z_{obt}

     d.) What is z_{cv}? 1.645 (p. 201)

(a) Should H_o be rejected? What should the research conclude?

The researcher would fail to reject Ho. The absolute value for the calculated Z score (-1.59) is less than the absolute value for the critical z-score value (z_{cv} = -1.645). Because 1.59 is less than 1.645, we fail to reject the null hypothesis. The researcher can conclude that the sample toothpaste did not significantly reduce the number of cavities when compared to other current toothpaste brands used in the general population.

(b) Determine the 95% confidence interval for the population mean, based on the sample mean.

$CI = \bar{X} \pm z(\sigma_{\bar{X}})$

$CI_{(95\%)} = \bar{X} \pm z(\sigma_{\bar{X}})$

$CI_{(95\%)}$ = sample mean $\pm$ z score(standard error of the mean)

$CI_{(95\%)} = 1.5 \pm (1.96)(0.145)$

$CI_{(95\%)} = 1.5 \pm (.2842)$

$CI_{(95\%)} = 1.2158 - 1.7842$

- Question 4. Henry preformed a two-tailed test for an experiment in which N = 24. He could not find his table of t critical values, but he remembered the t_{cv} at df = 13. He decided to compare his t_{obt} with this t_{cv}. Is he more likely to make a Type I or a Type II error in this situation?

t_{cv}: ±2.1604 at df = 13
df = 24 – 1 = 23
t_{obt}: ±2.069 at df = 23

If you look at the critical t-table in the appendix, you will see that the critical values decrease as df increase. Thus, with larger df, you will have a lower critical t-value. Because we are reading the critical t-value for df of 13 instead of 23 (for the sample), then we are using a higher critical t-value to compare to our observed value. Thus, we are less likely to reject the null hypothesis with a smaller degrees of freedom. As such, we are more likely to make a type II error, failing to reject the null hypothesis when it is false.

Henry is more likely to make a Type II error. A type II error occurs when the researcher fails to reject Ho but in it should have been rejected. As the degrees of freedom increase, the critical values decrease. He is using a critical value based on a df=13 which is 2.093. He should have used df=23 which is 2.069. He is more likely to fail to reject Ho because of the larger critical value.

- Question 6. A researcher hypothesizes that individuals who listen to classical music will score differently form [1]the general population on a test of spatial ability. On a standardized test of spatial ability, μ = 58. A random sample of 14 individuals who listen to classical music is given the same test. Their scores on the test are 52, 59, 63, 65, 58, 55, 62, 63, 53, 59, 57, 61, 60, 59.

a.) Is this a one- or two-tailed test? Two-tailed test.

b.) What are H_0 and H_a for this study? H_0: μ = 58 and H_a: μ ≠ 58

c.) Compute t_{obt}. Example: $t_{obt} = (\bar{x} - \mu) * ((\sqrt{n}) \div sd)$

Standard Deviation = Find the mean. Then for each number: subtract the mean and square the result. Then work out the mean of those squared differences. Finally, find the square root of that number.

$t = (46 - 58) * ((\sqrt{14}) \div 1.252)$
$t = -12 * (3.74 \div 1.252)$
$t = -12 * 2.987$
$t_{obt} = -35.844$

d.) What is t_{cv}?

df = 14 – 1 = 13
t_{cv}: ±2.1604 at df = 13

e.) Should H_o be rejected? Yes, because t_{obt} = -35.844 is less than 13. [2]What should the research conclude? The researcher should conclude that H_0 will be rejected and it will likely [3]be a Type I error.

f.) Determine the 95% confidence interval for the population mean, based on the sample mean. Example: $CI = (\bar{x} \pm t_{cv}) * (\sigma \div ((\sqrt{n})$

CI = ±0.66 and the range for the true population mean: 45.34 to 46.66
$CI_{(95\%)}$ = +56.805 (lower limit) - + 61.19456 (upper limit)

- Question 8. A researcher believes that the percentage of people who exercise in California is greater than the national exercise rate. The national rate is 20%. The researcher gathers a random sample of 120 individuals who live in California and finds that the number who exercise regularly is 31 out of 120.

a.) What is x^2_{obt}? $2.55 = x^2_{obt}$
b.) What is df for this test? One degree of freedom.
c.) What is x^2_{cv}? $3.841 = x^2_{cv}$
d.) What conclusion should be drawn from these results?
H_0: Cali Exercise ≤ NA Rate or $\mu_1 \leq \mu_2$
H_a: Cali Exercise > NA Rate or $\mu_1 > \mu_2$

2.55<3.841 therefore, H_0 should be rejected because the cv is greater than the obtained score and therefore, there is a significant difference between the variables.
You reject the null hypothesis when the calculated value is greater than the critical value. You do not reject the null hypothesis if the calculated value is less than the critical value. Here, you see that the calculated value of 2.55 is less than the critical value of 3.841; therefore, we do not reject the null hypothesis.

Jackson even-numbered Chapter Exercises (pp. 273-275).

- Question 2. A student is interested in whether students who study with music playing devote as much attention to their studies as do students who study under quiet conditions (he believes that studying under quiet conditions leads to better attention). He randomly assigns participants to either the music or no-music condition and has them read and study the same passage of information for the same amount of time. Subjects are given the same 10-item test on the material. Their scores appear next. Scores on the test represent interval-ratio data and are normally distributed.

WITH MUSIC	WITHOUT MUSIC
6	10
5	9
6	7

5	7
6	6
6	6
7	8
8	6
5	9

a.) What statistical test should be used analyze these data?
Two-sample t Test

b.) Identify H_0 and H_a for this study?
H_0: Study w/Music ≥ Study w/out Music or $\mu_1 \geq \mu_2$
H_a: Study w/Music < Study w/ out Music or $\mu_1 < \mu_2$

c.) Conduct the appropriate analysis.

$$t = \frac{\bar{x_1} - \bar{x_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \qquad S = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$$

$\bar{x}_1 = 6$ and $\bar{x}_2 = 7.56$
$n_1 = 9$ and $n_2 = 9$
σ or $s_1 = 1$ and $s_2 = 1.4675$
$n-1 = 8$
$t_{obt} = -2.635$
$t_{cv} = 1.746$

$t_{obt} = \frac{6-7.56}{.60311}$

$t_{obt} = \frac{-1.56}{.60311} = -2.5866$

d.) Should H_0 be rejected? Yes. What should the researcher conclude?
H_a: Study w/Music < Study w/ out Music or $\mu_1 < \mu_2$
What is the critical value?

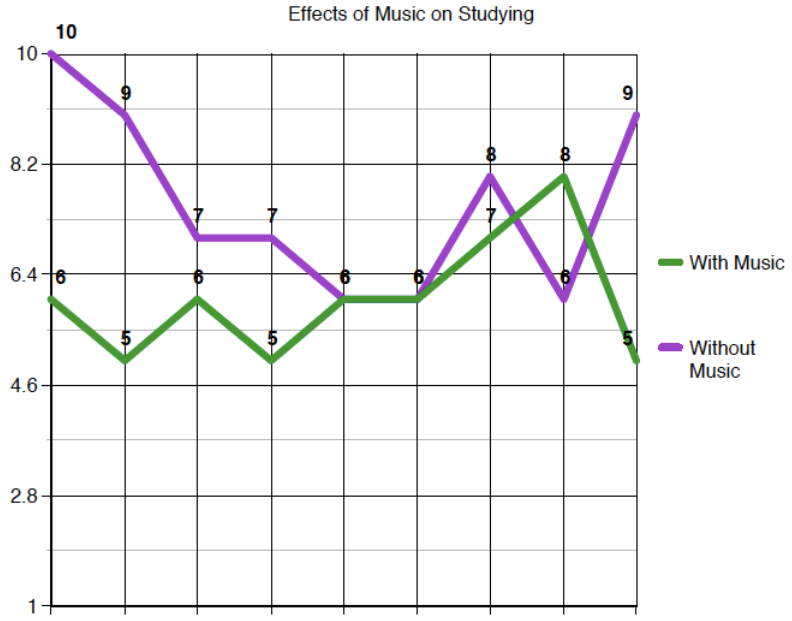
e.) If significant, compute and interpret the effect size.
d = 1.24 and r = -0.53

$r^2 = (t^2) / (t^2 + df) = (-2.488)^2 / ((-2.488)^2 + 16) = 6.190 / 22.190 = 0.2790$

The effect size of 0.2790 is large.

$2.5776^2/2.5776+16= 6.644/22.644= .29$

f.) If significant, draw a graph representing the data.



g.) Determine the 95% confidence interval.

H_a: Study w/Music < Study w/ out Music or $\mu_1 < \mu_2$

-2.38 to -.74

- Question 4. The researcher in exercise 2 decides to conduct the same study using a within-participants design to control for differences in cognitive ability. He selects a random sample of subjects and has them study different material of equal difficulty in both the music and no-music conditions. The study is completely counterbalanced to control for order effects. The data appear next. As before, they are measured on an interval-ratio scale and are normally distributed; he believes that studying under quiet conditions will lead to better performance.

WITH MUSIC	WITHOUT MUSIC
7	7
6	8
5	7
6	7
8	9
8	8

a.) What statistical test should be used analyze these data?

Within-Subjects design using a correlated groups t Test.

b.) Identify H_0 and H_a for this study?

H_0: Study w/Music ≥ Study w/out Music or $\mu_1 \geq \mu_2$

H_a: Study w/Music < Study w/ out Music or $\mu_1 < \mu_2$

c.) Conduct the appropriate analysis.

Example: $t = \bar{d} - 0 \div {}^{s}\bar{d}$

$t = .05 - 0 = .05$

$t = .05 \div .5949$

$t_{obt} = .084$

$t_{cv} = 2.015$

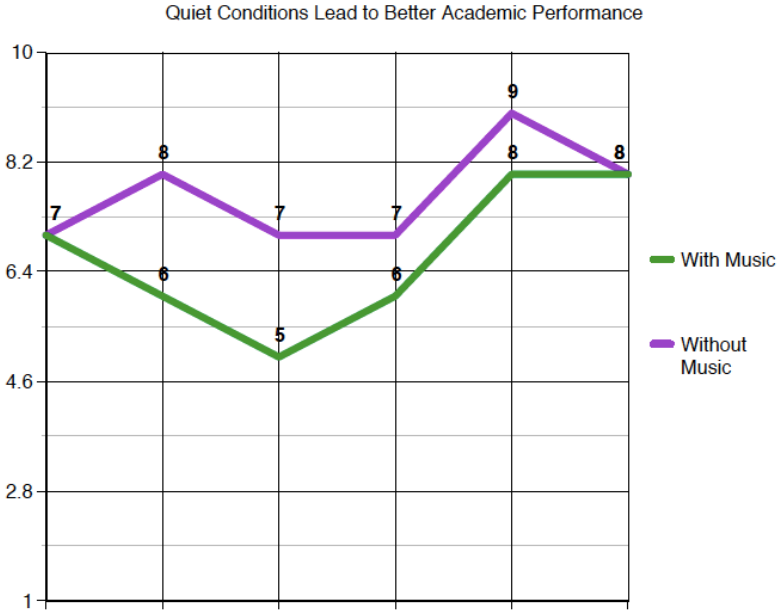
d.) Should H_0 be rejected? Yes. What should the researcher conclude?

H_a: Study w/Music < Study w/ out Music or $\mu_1 < \mu_2$

e.) If significant, compute and interpret the effect size.

d = 0.075 and r = 0.038

f.) If significant, draw a graph representing the data.



g.) Determine the 95% confidence interval.

-2.89 to 2.99

- Question 6. Researchers at a food company are interested in how a new spaghetti sauce made from green tomatoes (and green in color) will compare to their traditional red spaghetti sauce. They are worried that the green color will adversely affect the tastiness scores. They randomly assign subjects to either the green or red sauce condition. Participants indicate the tastiness of the sauce on a 10-point scale. Tastiness scores tend to be skewed. The scores follow.

RED SAUCE	GREEN SAUCE
7	4
6	5
9	6
10	8
6	7
7	6
8	9

a.) What statistical test should be used analyze these data?
Wilcoxon Rank-Sum Test

b.) Identify H_0 and H_a for this study?
H_0: $Md_{red} = Md_{green}$
H_a: $Md_{red} > Md_{green}$

c.) Conduct the appropriate analysis.
Data are skewed and I have assumed these are two independent samples (that is the subjects tested are not the same in both groups).

$n_1 = 5$ and $n_2 = 5$
$N = 7$
$t = 28$
$w_{cv} = 19$

d.) Should H_0 be rejected? No. What should the researcher conclude?
H_0: $Md_{red} = Md_{green}$ The result is not significant and therefore, the null hypothesis should fail to be rejected. There appears to be no significant difference in tastiness between the green and red sauce.

- You notice in your introductory psychology class that more women tend to sit up front, and more men sit in the back. To determine whether this difference is significant, you collect data on the seating preferences for the students in your class. The data follow.

	MEN	WOMEN	Total
Front of the Room	15	27	42
Back of the Room	32	19	51
Total	47	46	93

a.) What is x^2_{obt}? $x^2 = 6.7323$

b.) What is df for this test? 1

c.) What is x^2_{cv}? 3.841

d.) What conclusion should be drawn from these results? We do not reject the null hypothesis because the critical value does not exceed the obtained score. With regard to sitting in the front or back of the class, there is an association between men and women. That is, more women tend to sit up front, and more men sit in the back. However, there is no statistical difference.

2. What are degrees of freedom? Degrees of freedom are the number of samples in a group that a researcher is able to change (Jackson, 2012). How are they calculated? To find the degrees of freedom take n (the total number of samples) minus one (1).

3. What do inferential statistics allow you to infer? Inferential statistics include procedures for drawing conclusions that are based on data collected during an experiment. These conclusions involve making inferences about the population (Jackson, 2012).

4. What is the General Linear Model (GLM)? The General Linear Model (GLM) incorporates several statistical analyses used in applied and social research (Trochim, 2006). Why does it matter? The GLM matters because it is the foundation for the t-Test. In this case hypothesis testing uses the general linear model, which can be designed in two ways: independent testing or within-group testing (Jackson, 2012; Trochim, 2006).

5. Compare and contrast parametric and nonparametric statistics. Why and in what types of cases would you use one over the other? Nonparametric statistical methods are less reliable because a nonparametric correlation only includes ordinal position of pairs of scores while parametric correlation uses information about the mean and deviation from the mean (Schmidt, 2010). Parametric statistics includes interval and ratio scales and nonparametric statistics includes nominal or ordinal scales. Parametric statistics would be used when comparing performance of two different groups to determine whether they represent the same population or different populations (Jackson, 2012). Nonparametric statistics would be used to identify differences in ranks on a variable between groups (Jackson, 2012).

Fig. 1 (Jackson, 2012)

Parametric Statistics	**Nonparametric Statistics**
Is a test for a two-group between subjects design or matched subject design for independent or correlated groups.	Is a test for a two group between subjects design or matched subject design.
Compares performance of the groups to determine whether they represent the same population or different populations.	Identifies differences in ranks on a variable between groups.
Analyzes whether each individual performed in a similar or different manner across conditions.	Will identify difference in signed ranks on a variable for correlated groups.

6. Why is it important to pay attention to the assumptions of the statistical test? It is important to pay close attention to the assumptions of statistical tests because the

assumptions are the foundations of any research study. If your assumptions are found to be invalid then the researcher must conclude that the test is flawed (Trochim & Donnelly, 2008). What are your options if your dependent variable scores are not normally distributed? According to Laerd Statistics (2013) "the t-test is described as a robust test with respect to the assumption of normality, (section 5, para. 1). This means the t-test can still be used because deviations away from normality do not have a significant influence on obtained scores. There is one exception, and that is if the difference in the size of the groups is greater than 1.5, largest to smallest (Laerd, 2013).

References

Jackson, S. L. (2012). *Research methods and statistics: A critical thinking approach*. Belmont, CA: Wadsworth Cengage Learning.

Laerd Statistics. (2013). Independent T-Test - An introduction to when to use this test. Retrieved from https://statistics.laerd.com/statistical-guides/independent-t-test-statistical-guide.php

Schmidt, F. (2010). Detecting and correcting the lies that data tell. Retrieved from http://pps.sagepub.com.proxy1.ncu.edu/content/5/3/233.full.pdf%20html

Trochim, William M. (October 20, 2006). The Research Methods Knowledge Base, 2nd Edition. Retrieved from http://www.socialresearchmethods.net/kb

Trochim, W. M., & Donnelly, J. P. (2008). *Research methods knowledge base*. Mason, OH: Atomic Dog/Cengage Learning.